

# EPIWORK WP3 FOURTH YEAR INTERIM REPORT

## 1. Publishable summary

### *Project objectives for the period*

*Please provide an overview of the project objectives for the reporting period in question, as included in Annex I of the Grant Agreement.*

*Please include a summary of the recommendations from the previous reviews (if any) and indicate how these have been taken into account.*

*Please provide an overview of the project objectives for the reporting period in question, as included in Annex I of the Grant Agreement.*

*Please include a summary of the recommendations from the previous reviews (if any) and indicate how these have been taken into account.*

WP3 is concerned with the development of the Epidemic Marketplace, <http://epimarketplace.net>, an information platform to mediate access to distributed collections of public health data, offering an easy and safe way to share data for those data providers who want to collaborate with epidemiological modellers. Researchers will use this platform in multiple ways: i) as catalogue of data sources containing the metadata describing existing databases; ii) as a forum to publish information about their own data, seeking modellers to collaborate with, and/or to seek sources of data that could be of interest to their epidemiological modelling efforts; and finally, iii) as the host of mediating software that can automatically process queries for epidemiological data available from the information sources connected to the platform.

In the first years, the development was focused on the deployment of the infrastructure for the platform, characterizing epidemic data and developing a common meta-model for querying data of interest for epidemiological modelling. The platform mediates access to distributed collections of public health data, offering an easy and safe way to share data for those data providers who want to collaborate with epidemiological modellers through a publicly available Web interface. Development is now focused on a new version of the Epidemic Marketplace with much improved usability for curating epidemics data, together with an improved access control framework. We continue researching ontologies for interlinking epidemic data using semantic web standards and working on methods for assisting collaborators in easily find each other, and define mutual voluntary agreements for sharing their data, along with intensified work on populating the Epidemic Marketplace with epidemic data using the information collection tools and services developed by the consortium in previous years.

### **2.1 Background Information on WP3**

Work package number	3	Start date or starting event:	Month 1
---------------------	---	-------------------------------	---------

Work package title	Information platform							
Activity type	RTD							
Participant number	12	1	2	3	4	5	9	10
Participant short name	FFCUL	ISI	FGC-IGC	TAU	MPI-DS	AIBV	BIU	FBK-IRST
Person-months per participant	82	60	4	8	42	2	11	6

This Work Package is lead by FFCUL. The whole WP3 activity is structured into four tasks:

### Task 3.1 – Data Collection.

**Participants:** FFCUL, ISI, FBK-IRST, BIU, MPI-DS, FGC-IGC, AIBV.

**Description:** Realistic simulations of epidemic processes crucially depend on the availability of datasets describing human behaviour and pathogen-host interactions. Datasets include population movement data, social and behavioural data, health related data, geographic data, detailed geo-temporal epidemic incidence and immunization data, pathogen evolution and multi-strains circulation data. Data can come from a variety of different sources, including hospital records, country statistics, Web content, and others. It can range from a global scale, such as the worldwide air transportation infrastructure, down to the detailed description of individual **activities** at a minute-by-minute scale. This task will create a catalogue of databases of epidemiological data across Europe, with extensive meta-data describing the main characteristics of the available information sources. This catalogue will be integrated with a collaborative platform that will be set up for online discussion and exchange of meta-data among the participants.

### Task 3.2 – Meta-Model Design.

**Participants:** FFCUL, IGC, ISI.

**Description:** While some of the previously mentioned datasets are freely available on the Web (e.g. WHO Global Health Atlas, Eurostat), they are often scattered in different repositories, cover partial regions of the world and come in different formats, according to different standards and classifications. The project envisions a unified and integrated approach for the management of these resources, with the design and implementation of an Epidemic Marketplace Platform, publicly available on the web. The platform supports the sharing and management of epidemic datasets and resources as well as their rating, annotation, and selection. It is an on-line social networking site that will serve researchers,

practitioners, and educators all over the world to foster a virtual community for epidemic research. It will support the exchange of resources as well as user interactions. Based on a Web2.0 approach, users will become active participants, sharing information and data, and collaborating online, rather than being satisfied with a passive information consumer/viewer role. We envision proposing a simple reference format, which will facilitate the navigation and use of the datasets. Each dataset will come with a metadata file, signalling general metadata for resource management, containing data such as: the title, the date of submission, version, the source of the data and coverage. Moreover, the metadata will include information for a more thorough description of the data included in the dataset, providing a framework for a more specific description, for example, of epidemiologic and geographic data,. The Marketplace will support flexible and intuitive tools for navigation and selection of resources. Standard classifications as well as tagging systems proposed by users will be supported.

### **Task 3.3 – Epidemic Marketplace Platform.**

**Participants:** FFCUL, ISI.

**Description:** This task will implement a platform based on the integration of grid technology and publicly available services and software on the web to support the sharing and management of epidemic datasets and resources as well as their rating, annotation, and selection. The Epidemic Marketplace Platform will be an on-line social networking site that will serve researchers, practitioners, and educators all over the world to foster a virtual community for epidemic research. It will support the exchange of resources as well as user interactions. Based on some of the Web2.0 characteristics, users will become active participants, generating information and providing data for sharing, and collaborating online, rather than being satisfied with a passive information consumer/viewer role. More specifically, researchers can use and contribute to the Marketplace in several different ways. They can: (1) use it as a catalogue of data sources containing the metadata describing existing databases; (2) view, download, tag, and comment on the available resources; (3) provide compliant datasets and relevant information; (4) use it as a forum where to publish information about their own data, seek modellers to collaborate with, share and distribute their new findings.

### **Task 3.4 – Evaluation and monitoring of the use of the catalogue and collaboration services.**

**Participant:** FFCUL.

**Description:** This task involves the monitoring of epidemiological data exchanges performed through the mediating services platform. The evaluation will assess not only the coverage of the catalogued resources, but the users' satisfaction with the user interface and

integrated collaborative tools made available through the epidemiological marketplace platform. More importantly, the analysis of the collected datasets and their annotations and usage will provide a rich environment for deriving an epidemiologic ontology, which will help further on the integration and communication among the community of epidemiologists.

## 2.WP Work progress and achievements during the period

- *A summary of progress towards objectives and details for each task;*
- *Highlight clearly significant results;*
- *Team publications within the cope of the project (please provide a pdf of the publication if possible);*
- *List of outreach activities (conferences, Invited talks, presentations, workshops, tutorials,);*
- *List of press releases or media coverage, any particular dissemination activity;*
- *unanticipated finding, opportunity etc.*
- *If applicable, explain the reasons for deviations from Annex I and their impact on other tasks as well as on available resources and planning;*
- *If applicable, explain the reasons for failing to achieve critical objectives and/or not being on schedule and explain the impact on other tasks as well as on available resources and planning (the explanations should be coherent with the declaration by the project coordinator) ;*
- *a statement on the use of resources, in particular highlighting and explaining deviations between actual and planned person-months per work package and per beneficiary in Annex I (Description of Work)*
- *If applicable, propose corrective actions.*

## 2.2 *Progress in the Reporting Period*

*[please refer to the progress report by FFCUL, the WP3 leader, on this WP]*

## 2.3 *Outcast for WP3*

Focus for the second semester of the fourth year will be on providing the methods to further achieve INTEGRATION with the EPIWORK partner's systems and tools. In addition, we will be focusing on interlinking epidemic data in the Epidemic Marketplace using ontology engineering technologies. In parallel, we will continue developing the software of the Epidemic Marketplace, improving its interfaces and populating it with epidemic datasets. Much attention will be given to the design and development of a new front-end with a new user interface. The one that we currently have online is no longer satisfactory, to a large extent because the user interactions and the concepts have been so extensively reviewed in the past two years.

The envisaged activities will mainly involve:

- Redesigning the graphic user interface of the EM.
- Continuing work on populating and documenting the EM Repository.
- Addressing access control for data sharing within the community of epidemiologists.
- Providing an improved version of the curator's interface with support for user-created and managed collections.
- Semi-automatic annotation of resources using the NERO ontologies.
- Extend and improve the NERO Ontologies.
- Evaluation of the platform.
- Dissemination: working, among others, on a journal paper that will explain to the community of epidemiologists the benefits of ontologies for curating data and showcase the Epidemic Marketplace as a leading demonstrator of the application of these technologies to epidemics forecasting. New publications showcasing other aspects of the design of the Epiwork platform are also planned.

To complete these tasks we are in need of a confirmation that the project will have a 6-months extension at no extra costs. We have been working under the assumption that such extension would in fact be granted.